The function of the semiotic principle in establishing the claims of a pseudo or proto-science (graphology) to the status of empirical science

MARGARET GULLAN-WHUR

The claims of graphology: Its face validity and some ground-clearing

Graphology is the study of the interpretation of handwriting. Its first claim is that specific movements in handwriting (henceforth hw) indicate specific personality traits. Secondly, it claims to be better at cryptanalysis, interpreting hidden signs, than any diagnostic tool available to empirical psychology. Thirdly, it claims to be a science. It has no cause for complacency. Frequently accused of failing the least rigorous of scientific criteria (i.e., that tests show some unexplained correlation between hw movements and personality traits), graphologists have not yet demonstrated their technique's validity to the satisfaction of skeptical scientific inquirers. Says Furnham (1988): 'Some, albeit few, studies show extra-chance results linking hw to such personality traits as neuroticism. ... From any objective and dispassionate evaluation ... graphology is quite simply invalid'. Cox and Tapsell (1991): 'At a holistic, interpretative and descriptive level, graphological reports can provide some valid information, but insufficient to recommend graphology as a single diagnostic tool'. Fowler (1991): 'Predictive validity coefficients in the range 0.1 to 0.3 [on a scale 0.0–1.0] seem a more accurate summary of research evidence than the zero quoted by The British Psychological Society'. Discounting for the moment tests which show far greater validity, these conclusions are neatly summarized by Goldberg (1986): 'Although the link between hw and personality is probably very weak, it is not zilch'.

This collective verdict leaves unclear the status of graphology. Pseudo-science or proto-(primitive) science? Misconceptions abound even among those who attempt to show its face validity. In the first part of this paper I shall clarify graphology's claims. In the second part I shall suggest that the spotlight of semiotic inquiry illuminates graphology's place as an ethological tool within the current practice of natural science, and acts as a beacon for improved research methodology.

Bar-Hillel and Ben-Shakhar (1986) offer nine apparent strengths of
graphology as a device for personality assessment. While emphasizing that these *a priori* claims are no substitute for empirical investigation, the authors suggest the following advantages of hw analysis: (i) It relies on self-generated, self-expressive behavior samples. (ii) Hw exhibits different and shared personality structures. (iii) Hw is rich in features. (iv) Hw is a unique, stable feature of an individual even under different conditions. (v) Hw is within the ability of everyone undergoing a personality test. (vi) Hw is cheap and easy to acquire as a sample for analysis. (vii) Hw is available without consent or cooperation from the contributor. (viii) Hw is not amenable to alteration or feigning. (ix) Hw is said to be capable of divulging all the features of personality, including traits such as honesty, for which no other tests exist.

Almost all these assumptions are attacked by the authors above and in other critical literature. Graphologists themselves would deny or qualify iii, iv, v, vi, vii, viii, and ix. Hw is not always rich in features (iii). Both copybook hw, which shows no deviation from taught style, and calligraphic script, which is self-consciously styled, contain few features of interest to an analyst. Luria (1970) finds two distinct cortical systems to be involved in the production of hw. He analyzes hw disturbance in two cases of left parieto-occipital lesions where copying and slow writing based on optico-spatial analysis of letters was impossible, but quick writing based on automatized hw skill remained intact. Graphologists find only spontaneous hw rich in interesting features. Nor would they claim that hw is a unique, stable feature of an individual even under different conditions (iv). They object that hw is not and should not be entirely stable, that changes in hw over time reflect actual personality changes or growth. But tests have now decisively validated the stability, consistency, and individuality of script features for purposes of identification (Harvey 1934; Matwejoff 1934; Britt and Mensh 1943; Pophal and Dunker 1960; Fluckinger et al. 1961; Fischer 1962; Timm 1967; Ruenes 1967; Ullmann 1974; and Wallner 1975). Klimoski and Rafaeli (1983) concede that specific hw characteristics generally remain stable over time, and mention the studies of Timm, Fischer, and Prystav, whose correlation coefficients were all over 0.85, most over 0.9. Since a coefficient of 0.2 is taken to exceed chance, and 0.5 is considered to reflect clear validity, this aggregate score is high indeed. Note in the context of the later scope of this paper that the uniqueness and stability of hw traits were relatively quickly validated following post-evolutionary theory interest in overt expressive behavior, but also that the graphologists’ disclaimers touch vitally on the question of predicting future behavior from hw traits.

Writing continuous script may be within the ability of almost everyone undergoing a personality test (v), but the hw produced may not be
suitable for analysis due to problems connected with feature-richness (iii), stability or consistency (iv), disguise attempts (viii), or absence of the features sought (ix). For these same reasons, hw suitable for analysis may not be easy to acquire (vi). Graphologists maintain that samples flawed by copied or dictated material, lined paper, and uneven or moving surfaces will mitigate findings. Hw may be available without consent or cooperation from the contributor (vii), but because samples for mutual comparison must be written under similar conditions, a specially written passage is almost always required. Graphologists would admit that hw is to a limited extent amenable to alteration or feigning (viii), either in calligraphic script or other slow, stylized forms such as deliberate letter-disconnection. But it is established beyond question that not even a grasp of graphological principles could obliterate evidence of tampering (see Miller 1982 for details of eighty-seven tests). We can conclude that despite some potential for feigning, hw appears to offer substantial resistance to ‘test sophistication’, the process whereby candidates produce for the assessor the personal qualities they think are wanted. ‘Socially desirable qualities, such as ambition and team spirit, are relatively easy to fake on application and during interviews’ (Hubbard 1985). Assessment center questionnaires may also be vulnerable, whereas hw analysis is largely immune to hot tips on self-presentation. Finally, graphologists do not claim that hw is capable of divulging all the features of personality. Character traits sought may not be indicated in the hw, or they may be indicated in contradictory form (e.g., the callous murderer who is loving at home, or the embezzler who is ‘good with money’). Notoriously, graphologists refuse to guess age or sex, saying that only psychological sex is revealed in hw. Ben-Shakhar et al. (1986) accuse them of avoiding falsification of this overt and easily verifiable feature. Goldberg (1986: 285) comments, ‘If it turns out you can only [decide sex] with 70% accuracy, then you already know something about the likely predictability [sic] of more difficult, more subtle, and less important kinds of individual differences’. But actual age and sex do not express facets of personality. Hormonal conditions, on the other hand, are not necessarily restricted to anatomical, birth-registered sex, and may affect behavior (Carlson 1987: 473–477).

To sum up: despite the protests of graphologists, face validity is supported by high empirical validity with respect to (iv) stable features of individual hw even under different conditions; and (viii) impossibility of alteration or feigning. If not always ‘rich in features’, it usually contains several of interest to a graphologist (iii). Rarely, features are too sparse or the hw is insufficiently spontaneous for a particular analytic purpose (v). But it is relatively cheap and easy to acquire (vi), as those who
conduct tests invariably find. If hw is available without consent or cooperation from the contributor it may not be suitable as a sample for analysis (vii). It does not always divulge all the features of personality, but as initially claimed, it purports to interpret hidden signs in a technique otherwise unavailable to empirical psychology (ix). The validated assumptions of stability and individuality (iv) and impossibility of feigning (viii) would seem to make unequivocal the conclusion that hw analysis relies on self-generated, self-expressive behavior samples (i). We can add the further advantage that these samples are permanent pieces of behavior, to be reassessed over time.

The primary and insufficiently validated claim of graphology

That hw contains unique and stable patterns (a phenomenon on which the secure use of credit cards is based) in no way substantiates the a priori assumption that hw exhibits different and shared personality structures (ii), or that hw is capable of divulging any features of personality at all (ix). It is a further step to establish that the neurophysiological factors responsible for hw are linked with the stable underlying aspects of temperament which we call 'character', 'personality', or 'disposition' (Bar-Hillel and Ben-Shakhar 1986; Quesnel and Wade 1990). According to Naftali (1965), the neurophysiological factors influencing hw are: heredity (inborn movement tendencies); acquired movement patterns (partly directed by higher brain centers situated in the cortex, which controls conscious motor behavior); changing neuromuscular tension (including repetitive mannerisms, stiffness, flexibility, and coordination); and variable behavior patterns (such as nervousness or joy) which may be influenced by the purpose of the script. Naftali's fifth factor of conscious style creation also bothers Wing (1991: 13). But it has been shown above that with the exception of some non-spontaneous formations this is outweighed by unalterable individual features. Graphology's task is to validate a connection with that aspect of the hw process responsible for writing based on automatized hw skill.

It will be seen that our two residual a priori assumptions — that hw exhibits different and shared personality structures (ii) and that hw is capable of divulging all the features of personality (ix) — match graphology's claims as initially stated. The first, that specific movements in hw indicate specific character traits, is foundational to the study. Upon its validity depend the two secondary assumptions: that graphology excels in cryptanalysis, and that it is a science. To date this primary assumption hovers, in the eyes of most impartial observers, somewhere between Ben-
Shakhar’s ‘supreme article of faith’ (1986: 652) and a vexing, insufficiently validated hypothesis. The onus, say Bar-Hillel and Ben-Shakhar (1986), is on graphology to establish its feasibility and status as a scientific enterprise. To this end it must express a testable, repeatable, and falsifiable correlation between hw and character or personality. Though few scientists care to satisfy Popper’s supreme test of verisimilitude by setting out to falsify their hypotheses (Popper 1963: 36), they generally hold a self-critical attitude to be healthy. They consider their conclusions, being empirically derived and matters of contingent fact, to be open to correction. They simply believe their theory to be the most fruitful explanation currently available. Such modest claims are nevertheless made only after their hypothesis has shown itself capable of being tested.

It is objected (a) that graphology has not been tested, and (b) that it is not testable. Furnham (1988) maintains that ‘scientific scrutiny of graphology’s claims remains limited and may be on the decrease’, and Ben-Shakhar et al. (1986) cite a ‘paucity of serious research efforts’. In contrast, Quesnel (1990) condemns graphology on the basis of repeatedly insignificant test findings: ‘Researchers have made valiant attempts ... during the past seventy years ... to determine the existence of valid relationships between hw and a number of personality features’. I have personally collated two hundred research papers relating specifically to the validity and reliability of graphology (Gullan-Whur 1991). This bibliography largely excludes research carried out by graphologists themselves, but see Miller (1982) for two thousand mixed studies. It is probable that no other pseudo or proto-science has been subjected to such extensive testing. We must therefore conclude that the insignificance of many graphological test results is due (a) to graphology’s untestable assumptions or (b) to unstandardized and inadequate testing. Option (a) is a charge frequently made, and it requires and will receive in-depth investigation with the help of the semiotic principle. Option (b) is aired by Rafaeli (1983), who suggests that participant graphologists had not worked with the scales in the study; that the gestalt or global nature of graphology (interpreting a complex interaction of all the signs) was overlooked; and that the attribute sought was either not within the scope of graphological analysis or was not indicated in the hw. It is obviously unfair to test a technique on claims it does not make, for example that it can indicate the actual job or occupation the writer holds. If graphologists agree to participate in tests, it is up to them to insist on a straightforward hypothesis of their choice. I shall suggest that it is their general refusal or inability to do this, their willingness to speculate publicly on writers’ future behavior, and their powerlessness to capitalize on occasional highly successful test results which contribute to graphology’s poor standing.
Science demands that a correlation be shown to exist. It does not demand a causal story. Aspirin was used for many decades on the basis of instrumentalism, the procedure used by normal scientists when progress is cut off by ignorance of mechanisms. Aspirin eliminated pain: it 'worked', but no one could discover why or how. Although a causal explanation extends the usefulness of the correlation, it is not a necessary or sufficient criterion of a science. So Furnham's complaint that 'even if graphological analyses were valid, the theoretical basis of the method appears weak, non-explicit and non-parsimonious' (Furnham and Gunter 1987) and Ben-Shakhar's comment that 'if a correspondence were to be empirically found between graphological features and such traits as honesty ... it would be a major theoretical challenge to account for it' (1986: 652) are irrelevant to graphology's scientific status.

That I suggest that graphology may not halt its research at a pragmatic conclusion, but may ultimately demonstrate the causal mechanisms which produce the correlation it claims, is a measure of my belief in it as an embryonic or proto-science. Graphology does not need to be mysterious, transcending empirical enquiry, taken on faith like a religion. I believe it is capable of establishing itself as a science, but because it carries the stigma of pseudo-science it must submit itself to the severest of criticism and the most stringent of tests. In the following inquiry I shall scrutinize the remaining charges commonly made against graphology. A consensus of criticism in academic literature finds twelve further flaws in its methodology and application. (These criticisms will be referred to later as '1a', '1b', etc.)

**Group 1: Unsound methodology**

Graphological judgments are not testable, repeatable, falsifiable, or objective because: (a) Methods of analysis are rarely empirically based, or judgments are vague, generalized. (b) Reference to unobservable and therefore unverifiable personality features makes judgments unfalsifiable. (c) Inconsistent and unlimited personality variables are used. (d) Inconsistent and unlimited hw variables are used. (e) There is low inter-rater reliability (i.e., inconsistent judgments between graphologists about the same sample of hw) and occasional low intra-rater reliability (i.e., inconsistent assessments of the same hw sample by a single graphologist). (f) There is widespread use of subjective rather than objective judgment. (g) There is a faulty graphological view of what constitutes validation.
Group 2: As a projective technique graphology is weak or redundant

(a) Graphological judgments show up poorly when correlated with findings of psychometric tests. (b) Graphology is a redundant technique. (c) Its predictive capacity is low, especially relative to employment decisions.

Group 3: Graphology involves no special skill

(a) Graphologists just read the script content. (b) Graphological judgments are no better than chance. (b) Hw ‘analysis’ by non-graphologists is equally reliable.

Applying the semiotic principle

Semiotics is a methodological principle which has contributed to the transformation and reorientation of several disciplines (Sebeok 1975, 1976, 1977, 1978, 1979). It requires that aspiring sciences ask the kind of questions that the empirical sciences, preoccupied as they are with methodology and the problems of verification, will ask. It must be said that Sebeok recoils from what he calls ‘pseudo-scientific techniques’ which claim to receive signs from the environment (Sebeok 1976: 74). He does not, however, name graphology in his list of culprits. Since we have accepted as validated according to current scientific principles the a priori premise (i) that hw analysis relies on self-generated, self-expressive behavior samples, graphology may be thought to have cleared that hurdle, but this is not assumed. If no consistent correlation exists between hw traits and behavioral characteristics, then hw is no more a body language of interest to psychologists than is the imprint of a bottom on a deck chair.

I shall briefly outline the semiotic principle before submitting graphology to its rigours. In 1690 John Locke reapplied the word *semiotika*, previously used only for musical notation, to human communication, ‘the business whereof is to consider the nature of signs the mind makes use of for the understanding of things, and conveying its knowledge to others’ (Locke 1961: 385). The American philosopher and logician C. S. Peirce took up Locke’s scanty thesis and during the late nineteenth century turned it into a lifelong study. He held that ‘the entire universe is perfused with signs, if it is not exclusively composed of signs’ (CP 5.44), meaning that we only have our own perspectival experience and interpretation to acquaint us with reality. Peirce’s sign was ‘something
which stands for something to somebody in some respect or capacity'. He held that a science of semiosis (the relation of signifier to signified) provided the framework for the existence of any kind of human understanding, and that his thesis could be extensively reapplied. Subsequent developments have supported this vision. The work of Charles Morris and his pupil Thomas A. Sebeok has resurrected Peirce's thesis and reintroduced its broader applications to the behavior of men and animals, creating a separate line of study to that of semiology or linguistic semiotics, so called in deference to its exponent Ferdinand de Saussure. I place graphology on the Peirce-Morris-Sebeok line, though it should be noted that it is unique to the human species and is in some ways subject to the problems of language as a manmade phenomenon. But the transition from pre-hominid signal systems to human language must be described in terms of an organism's receptive information-processing system (Sayre 1976: 192). If we believe evolutionary brain processes to lie behind both language and writing processes, we cannot entirely separate anthroposemiotics (man's species-specific signalling systems) from zoosemiotics (the study of animal communication). A parallel may be drawn between the idiosyncratic marks and tracks left by individual animals, and the traits formed by writers whose learned technique of literacy has become mechanical and automatic. Graphology is well placed to construct testable hypotheses correlating character modes with it traits, since the 'track' is durable and can be precisely measured. It is no accident that graphologists speak of it traits: the French word means a groove or furrow; for nineteenth-century graphologists it meant the grooves scratched into paper by a quill pen. The question is whether we can correctly regard it as signalling patterns of behavior.

At least one behavioral psychologist admits that it is an attractive expressive field (Allport 1933). I will return the compliment by offering two attractions of methodological behaviorism. Despite its well-documented limitations, behaviorism's harshest critics admit that its principles of objectivity (Carlson 1987: 23) and the other features which enable it to perform useful functions today 'emerged from the animal laboratories where behaviorism itself was born' (Mackenzie 1977: 159). Graphologists would claim that significance depends not on a relatively sparse range of animal behavior, but on a complex interweaving of conscious and unconscious dispositional attitudes. Bar-Hillel and Ben-Shakhar (1986: 274) detect a graphological 'preference for predicting [sic] deep-lying unobservables to their observable correlates and perhaps even a disavowal of the relevance of behavioural criteria to the evaluation of their assessments'. While bearing in mind that the supreme value of graphology is claimed to lie in cryptanalysis, revealing hidden propensi-
ties, it is clear that for the initial purposes of testing and verification we have to have an outward mode of character. The verifiable significance of an assertion depends on recovering referents. Behaviorism's insistence that the only proper subject matter for scientific study in psychology is observable behavior allows us to distinguish between observable, verifiable (and possibly characteristic) behavior, and hidden 'dispositions' which may not be verifiable because they do not exist. The second attraction of this crude methodology is that it lets us dispense with the dubious term 'personality'. It frees us to speak of 'character modes', ways of behaving which or may not be related to some stable underlying disposition which would make them 'personality traits'. Our tendency to assume that behavior is monitored by a 'complex of psychological characteristics by virtue of which an individual is recognisably unique' (British Psychological Society) is considered by many to be an attributional error (Carlson 1987: Ch. 13.; Bar-Hillel and Ben-Shakhar 1986: 272; Johnson-Laird 1983: 475; Gregory 1984: 496). Graphology does insist that hw traits signify 'personality traits', but I shall question whether it is in its best interests to persevere with that further claim. Let us proceed with the simplified claim that hw 'tracks' signify character modes. There may be no such phenomenon as a 'personality', but there is undoubtedly behavior.

Semiosis can illustrate the potential significance of graphology. At its most basic level of interpretation the propositions on which knowledge of our universe depends are asserted through one of a triad of signs: icon, symbol, and index (Peirce 1866; Sebeok 1975: 90). The icon is something which functions as a sign by means of features of itself which resemble its object (or referent or signified). An example is the road sign for a hump-backed bridge. The icon is recognizable in certain graphological links between the hw trait or signifier, and the character mode or signified. An example is 'islands' in the script, where widely spaced words are said to reproduce visually the writer's need for social or psychological space. Iconicity plays an unscientific role in graphology because its semiosis is unverifiable. Peirce called it 'the sign of all possible lies' (CP 4.531). Sebeok warns that its contemplation sooner or later turns from legitimate semiotic concerns to intractable problems of identity, analogy, resemblance, and contrast (Sebeok 1979: 118). Iconic meaning involves both individual intention and social convention, so we can never be sure of the basis in meaning of the alleged resemblance. We shall return to the function of the icon with criticism Group 3 (Graphology involves no special skill).

The symbol is something which functions as a sign because of some rule of conventional or traditional association between itself and its
object, assigned by us with varying degrees of interpersonal agreement. Its semiosis is neither a necessary one nor one with a direct causal connection. Words are symbols, arbitrary signs chosen by us and made conventional signs through public adoption. Symbolic interpretation has greatly enriched the practice of graphology, and there is a wide debate on whether symbols and metaphors do reveal human truth. (Peirce classed metaphors as icons.) A distinction should be made between a closed system of metaphor, where definitions are made plain at the outset, and organic metaphors wherein the system can rampantly enlarge itself. Once the system reaches this point all interpretive control may be lost; the system subsides into an abyss where everything might be claimed to resemble everything else. It is clear that such ‘signalling’ is unsuited to an aspiring empirical science. But if a specified group of hw symbols is found to correlate scientifically with a specified set of character modes, we should suspect that the sign-system involved is not symbolic, but indexical.

The index is something which functions as a sign by virtue of some factual or causal connection with its object. Medical signs or symptoms are examples of such signalling. Signs are not entities but relationships, and the index is a relationship which exists within nature. It is discovered, not created. It differs from the above classes of sign in that semiosis depends on an actual causal relationship between natural, concrete objects. If graphology is to be an empirical science it has to show that it functions indexically, since only as an index is it susceptible to empirical verification. In order to be an index, the graphological sign must show itself to be discoverable, to exist within nature. Figure 1 shows how we can mark off scientific from non-scientific discourse, and how indexical criteria could be met by graphology.

Consider first an assumption inherited from the abbot Michon (1872): that certain hw traits indicate ‘nobility of spirit’. To claim that B is a sign of D (because D causes B) is to attempt to establish a correlation with a transcendental object or referent, and to assert an untestable correlation. (1a: Graphological judgments are not testable, repeatable, falsifiable or objective since methods of analysis are rarely empirically based, and 1b: Reference to unobservable and therefore unverifiable personality features makes judgments unfalsifiable.) Consider also in the light of 1a that it has been shown above that we do not, simply by establishing an AC correlation (mode of character indicates brain state) and a CB correlation (brain state causes hw), substantiate an AB correlation and validate our claim. We can apply to this problem the methodology of the now empirically recognized disciplines of zoosemiotics and anthroposemiotics. The semi-independent research group set up under
the auspices of the Max Planck Institute of Behavioral Psychology to investigate anthroposemiotics recognizes that the verification processes used on animals (e.g., the testing of an electrocommunication channel where several kinds of discharge alterations have been clearly shown to correlate with social situations) are ethically and practically unsuited to humans. It therefore concentrates on outward modes of behavior, those which make an appreciable impact on at least one of the interpreter’s sense organs. As in medicine, where indexical signs form the basis of diagnosis, specific events or behavioral modes are investigated, and the living brain is usually left alone. But human and animal sign processes are not sharply distinguished: the rhesus monkey’s simple stare signifying hostility can be equated with certain universal human gestures and facial expressions previously considered culture-bound (Seboek 1975: 90–94). Behaviorism’s perceptual skills have resulted in an admirable phenomenology of behavior. Identifying and discriminating minutely between responses in an ongoing stream of behavior may be the chief contribution of behaviorism to the future development of psychology (Mackenzie 1977: 169).

Graphology has an advantage over many ethological tools in being able to extend its inquiries beyond the conceptual level to the observation of behavior, and beyond that to the causal level. Neurological work on CA and CB correlations refutes Bar-Hillel’s and Ben-Shakhar’s claim
(1986: 270) that features of body movement are only loosely correlated with hw movement, if at all. (See the many studies recorded in Gullan-Whur 1991 which show validated dynamic relationships between observable mental states, neurophysical or biochemical states, and specific hw traits.) Following work by van der Gon and Thuring (1965) and Vredenbregt and Koster (1971) on electrical pulse mechanisms, Wing (1978, 1980) chose to investigate hw features connected with tension, release, and stress, factors which influence hw formations and relate to many testable graphological hypotheses. In 1991 he focused on departures from a taught model, and brought in a graphologist/psychologist to attempt correlations. Results showed statistically reliable links between ‘d’ base vertical height and unobtrusiveness, considerateness, independence, conformity, social control, and extroversion. He found that hw features were related to behavior, not to intelligence. Such testing of causal theories is very desirable, but it is the task of graphologists to implement the first stage of research by offering AB hypotheses which claim that B (a specified hw trait) is an indexical sign of A (a specified outward mode of character). The criticisms of Group 1 pinpoint the difficulties which stand in the way of such progress. They are all domestic disputes, the business of graphologists themselves. They concern the refining of the variables A and B to a scientific, testable status.

**Applying the semiotic principle: the A and B variables**

To resist 1c (Inconsistent and unlimited personality variables are used), referent A must be an agreed, standardly termed behavioral event or psychological description of a state of affairs. To satisfy 1d (Inconsistent and unlimited hw variables are used), B must be a specified, standardly termed hw trait or very limited group of traits. The unwillingness of any but the boldest of graphologists to assert a correlation containing the isolated and well-defined variables necessary for testing works directly against their study’s chances of achieving scientific status (Ben-Shakhar et al. 1986: 652; Furnham 1988: 66, 68). Let us look at the problem of isolating the variables A and B.

The nature of referent A (the character mode) is the most frequently disputed aspect of any sign system, but the success of the AB correlation hinges on it and on graphology’s ability to show that it exists. We are concerned here with sensory verification in relation to 1f (Use of subjective rather than objective judgment). The first difficulty lies in the need for the character mode to be evident, observable. Johnson-Laird distinguishes co-referential descriptions of (a) the mental model derived from
visual perception, which will be a single entity corresponding to a single, perhaps unique state of affairs; and (b) the mental model based on discourse, discussion, or theory (Johnson-Laird 1983: 407). The latter may take the form of a convergence of opinion based on value judgments. Here the identification of A as a fact requiring a theory or assumption does not exclude an indexical relationship, since, as with a limited symbol system, we claim that our sign is really affected by that object. The two kinds of mental models are sometimes difficult to distinguish, and the variable ‘intelligence’ is an example of one which might belong in either camp. It is strongly suggested that since subjects living in some African countries not only perform on average poorly in English intelligence tests, but also appear to doubt the intelligence of the examiner in posing his questions, such tests may be based on a larger component of cultural discourse than is generally admitted (Carlson 1987: 366). If we are to command agreement on the existence of our referent A, we should prefer, at this pre-scientific stage of graphology’s development, a mental model which derives from visual perception, and which is decided objectively rather than subjectively or intersubjectively. This decision has profound philosophical implications which will be discussed at the end of the paper, but I believe it is the right one.

A footnote on subjectivity: graphologists trained in or affiliated with teaching disciplines are repeatedly warned about it, and urged to restrict their judgments to the agreed correlations, which are the result of a century of graphological observation. They are also advised to watch out for subjectivity of interpretation through overemphasis on particular variables, a snare in normal scientific practice. But they are often asked to make blatantly subjective judgements. Variables 6, 7, and 9 in the Ben-Shakhar (1986a) study read: ‘V6: Overall quality of written essay; V7: aesthetic evaluation of script; V9: overall impression of writer’. There can be no objective measurement of these.

Personal attributes which commonly interest clients, such as honesty, intelligence, or loyalty, may have no observable referent. Consider ‘criminal tendencies’: this cannot constitute an outward, testable character mode. Firstly, a tendency cannot be evident, only a concrete state of affairs. A tendency is established from the results of studies which quantify actual character states. Secondly, what constitutes criminal behavior is relativistically decided; therefore no neural mechanism may ever be found to correspond with what we must now term ‘crime-committing’. A character mode which is visually apparent is not only essential for testing, but if the correlation is upheld, that mode can then in principle be tested for AC semiosis (character mode indicates brain state). Irrecoverable referents make a mockery of tests. The graphologist
employed by a Vancouver school board was asked to identify a possible paedophile in 10 hw samples. He found 9. In fact, there were none (BC Skeptics 1988). But he had been asked to look for a possible, not an actual paedophile. His findings could not be invalidated until all 10 sample writers were dead, and the true (meaning empirically observable) facts known about relevant instantiations of potentiality. Even then the full story about their ‘possible’ behavior would not be verified. That particular referent remains for ever irrecoverable. On the other hand, Sassi’s (1987) study, and Honel’s as re-analyzed by Nevo (1986), distinguish ‘criminal’ precisely and recoverably as the state of incarceration for a criminal offense. In Honel, the hw of 41 male prison inmates including larcenists, embezzlers, sex offenders, and murderers was compared with a control group of 48 men reported to be loyal, honest, responsible, and without criminal record. After a bias due to socio-economic factors (interesting in itself) was detected, a new sample set was created. Only a combined index based on the four graphologists’ reports discriminated in a statistical way between criminals and non-criminals. Without a breakdown of the hw traits on which their judgments depended, we cannot say if difficulties hinged purely on the esoteric nature of the personality variable ‘criminality’, or in part on low inter-rater reliability (1f).

Low inter-rater reliability is inextricably linked to the problem of isolating and choosing variables. Low intra-rater reliability, on the other hand, is simply due to poor graphological training or a lack of understanding of the need for standard correlations. Peeples (1990) and Williams (1989) point out that low intra-rater reliability in some graphologists pulls down reliability as a whole. Peeples analyzes this in terms of training, noting the many different schools of graphology. But until there is agreement on essentials, test results will not accurately reflect the potential of graphology. Without inter-rater reliability there can be no validity. Nor can there be a ‘profession’ of graphology. Professional bodies may condone disagreement between their members over details, but if basic assumptions crumble, so does professional practice. The examinations set by graphological organizations should iron out this problem, but unfortunately there is not yet inter-organizational agreement. Gullan-Whur (1986) sought to unify continental and American hw trait significance.

Reliability is equally tied to the second difficulty with the A referent: namely, the refusal to limit variables (1c). Graphologists consider the use of single variables to be bad practice, since it oversimplifies character modes. To be sure, scientific referents are not always single or simple: e.g., pink spots on the skin may indicate conditions v, w, x, or y, and
next year the hypothesis may be revised to include z. Bearing in mind
later testing of CB and CA correlations, we should not make the mistake
of imagining that any psychological mode can be more complicated than
the brain states we believe to give rise to it. A sign must allow itself to
be clearly read; so why, at this pre-scientific and vulnerable stage, put
forward complicated hypotheses which may need immediate revision even
if they survive testing, when graphology could gain overnight credibility
through the corroboration of a clear and bold one? The price of
twentieth-century scientific status is the demonstration of straightforward
correlations in interesting areas of psychology.

The third difficulty connected with the referent A is that of the
psychological description of the character mode or trait. Descriptions
of psychological phenomena can only with great effort be expressed through
a uniform discourse, but standard terminology is vital if the referent is
to be indisputably recovered. Propositional representations which have
no overlap in their arguments can, in fact, be co-referential: Johnson-
Klimoski (1983: 195) found that ‘Low intercorrelations, interpreted as
indications of low reliability, may actually reflect different interpretations
of the same constructs’. The abbot’s ‘nobility of spirit’ could be, in fact,
the same character mode as ‘integrity’ or ‘courage’, but if psychological
descriptions are not standardized the chances of successful testing are
reduced. This can be as much the fault of the researchers as the grapholog-
gists. Both parties must ensure that all definitions and degrees of defini-
tions are settled before the test takes place, or no significance can be
established. A vague term such as ‘intelligence’ or ‘aggression’ must be
subject to agreed criteria for testing purposes, or what seems to be an
agreed psychological description may prove useless as a means of recovering
the referent.

We turn now to the problems of the B signifier or expression, the hw
trait. Many graphologists would reject any sign embodying a solitary hw
trait, and long lists of possible or interchangeable signifiers are given.
But if no one variable is considered necessary or sufficient, definitions
are out of the question and testing becomes an impossibility (1d:
Inconsistent and unlimited hw variables). For testing purposes we require
single or limited variables evident to any eye. I am unsympathetic to the
view that no hw traits can be autonomously tested, though some are
more easily isolated than others. In 1970 Wallner chose for testing hw
variables in certain clearly defined areas. These were added selectively to
a test battery which was already yielding good results (Wallner 1975).
Quesnel (1990: 24) notes that such atomistic features of hw as slant, size,
and spacing are the most reliably assessable and measurable; but in most graphological practice all these have multiple interpretations.

At this point many graphologists decide that scientific (testable) graphology is bad graphology, while good graphology, full of complex synthesis, is poor science. Others offer as a compromise the gestalt, the concept of overall or guiding images which provides for hypothesis single variables of complex origin — that is, generalized behavior modes signified by equally diffused patterns of signs in hw. Sadly, such variables are seldom amenable to testing by impartial observers. A graphologist would need to stand by to assess ‘disturbed rhythm’ in a script, and so would be subjectively instrumental in the judgment while the testers did not perceive the sign. Crumbaugh (1986; Crumbaugh and Stockholm 1986) claims that only global or holistic validation is possible for graphology, that the meaning of each sign changes slightly as it interacts with the others. If this is so, scientific validation of graphology seems unlikely indeed. It is not clear how Crumbaugh’s spontaneous ‘personality sketch’ of a friend, written in undefined psychological terms, can ‘match’ a single-page summary of hw trait analysis written independently by a graphologist. Problems of incompatible description, or attributes emphasized in one while missing in the other, loom large. Williams (1989) correlates informal descriptions with hw analyses, but only after both graphologists and friends have been allocated agreed definitions of a limited number of personality traits. Note that, despite this, Williams’s correlation coefficients are reduced by unreliability, since the graphologists did not have to analyze in terms of specific hw traits.

Applying the semiotic principle: the AB correlation

So we come to the sign which represents graphology’s central claim: that B (the signifier or hw trait) indicates A (an outward character mode or state of affairs). To satisfy scientific criteria this correlation must be accessible to any adequate observer by means of non-perspectival variables. I suggest two further provisos. (1) Before AB hypothesis testing, trait norm studies using large numbers should be carried out, and mean curves for the population as a whole constructed. This makes assessment of deviation or distinctiveness in behavior more accurate: for example, a score of 2 on a scale of 0–10 cannot be taken at its face value, since it may represent well above average endowment in that particular trait (Ness and Coleman 1991). Wing finds average levels for all the hw traits he investigates before making experimental tests. Williams (1990) demonstrates an average spread for each personality variable. (2) An adequate
number of samples should be used; a handful is insufficient for a fair test. Results show that correlations which are insignificant over a very small sample may become pronounced when larger numbers are analyzed (Williams 1989).

Let us consider two validated hypotheses. The testers were impartial in that they were non-graphologists and had no vested interest in the corroboration of the hypotheses. (It is noteworthy that few tests carried out by graphologists reach disappointing conclusions, though well-conducted tests have produced correlations which may surprise — e.g., Griffiths 1991.) In these tests the referents or signifieds were states of affairs, namely those of being medical doctors (Goldsmith 1976) and being women of established education or intelligence (Epstein et al. 1961: 385–392). The first of these testable hypotheses was that the hw of medical doctors is less legible than that of other individuals. It was upheld. The second study concluded from its observations that ‘the female adult who continues to write in the fashion that would have pleased her elementary school teacher is less likely to be as well-educated, as bright or as mature as the adult who has worked her way out of the school-copy rut and has evolved a more efficient and original way of writing’. Correlation coefficients varied according to the hw trait tested, but were overall significant. For the first study a control group of non-doctors (randomness of sample unknown) was required to give samples. For the second, best approximations of the median age, I.Q., and completed grade of 150 females registered with a Testing and Advice Center were calculated and used as the cut-off point. In these tests six Group 1 criticisms were countered: (a) method of analysis not empirically based; judgments vague or generalized; (b) reference to unobservable and therefore non-verifiable personality features; (c/d) use of inconsistent and unlimited personality and hw variables; (e) low inter-rater reliability; and (f) use of subjective rather than objective judgment. Critics may protest that legibility, even if observable, is to some extent subjectively decided, but Wing (1991) found judgments on legibility to be very consistent, producing an inter-rater correlation of 0.86. I have already commented on the intersubjective or cultural nature of intelligence tests: intelligence was assessed according to Wechsler scale intelligence quotients, whereas ‘well-educatedness’ was decided by observable education records.

The last criticism of the group (1g — there is a faulty graphological view of what constitutes validation) was not rebutted. Many graphologists would find the conclusions of these tests banal to the point of irrelevance. By and large, graphologists’ conceptions of validation form a considerable stumbling block to the credibility of their study. They do not want to offer bold hypotheses, since these give no impression of the synthesizing
skills needed to produce a personality profile or 'portrait'. Graphologists are taught explicitly to avoid stark correlations, and interrelate all personal and handwriting traits. Most do not intentionally rely on intuition, success stories, and unsupported hypotheses to validate their claims, but the great majority are untrained in the methods of exact or human science. When invited to answer Furnham's criticisms at a CSICOP conference, a leading graphologist 'tried to demonstrate that graphology ... has its roots in the same symbolism as does art, literature and psychology'. Another said that she put her faith in well-designed follow-up studies of clients' satisfaction levels (CSICOP 1988). Client satisfaction may get graphologists employment, but personal validation is not scientific validation. The 'Barnum effect' (Bar-Hillel and Ben-Shakhar 1986; Vestewig et al. 1976) convinces through the shock of flattery or generalized half-truths. Quesnel (1990) quotes Snyder and Larson (1976), where subjects lapped up fake personality descriptions. Lockowandt (1986) laments that the Pophal generation (c. 1940–60) neglected to train scientifically qualified successors, and that for this reason there are hardly any analysts suited to research. But there always have been scientifically trained graphologists (many eminent in scientific fields) who struggle to impress their colleagues with the need to consider scientific principles. Resistance, however, lingers. When the original version of this paper was read at Cambridge I was told by one delegate it should be burned, and a recent letter to a graphological journal claims: 'Scientists ... want graphology to be made "respectable" by imitating their jargon and technical vocabulary. Graphology works right under their eyes, and yet they try to prove that it doesn't' (The Graphologist 1991).

Antipathy to scientific method seems ingrained at a profound level in some graphologists: intransigence, not ignorance, predominates. I am reminded of the media coverage given to a positive report on palmistry (The Sunday Times, London, 18 June 1989). A group of pathologists at the Bristol Royal Infirmary, England, 'put palmistry to the test by comparing the length of life-line and age of death in 100 autopsies'. To the researchers' surprise, 'there was a trend. People who died young usually had shorter life-lines than those who lived to a great age'. Were palmists grateful for this scientific evidence handed to them as a gift? They were not. 'It would be extremely foolish', commented one, 'for anyone to expound the theory that the length of a person's life can be judged on the length of a line alone'. Did scruples over public consternation at finding short life lines provoke their disclaimers? No. The reaction was one of disdain at the oversimplification. 'If I must define briefly what a short life line means', one said, 'it's a trauma, the death of a "life". But I want to give all the alternatives'. However, most graphological organiz-
tions encourage or undertake rigorous research, and this attitude is becoming more widespread. Note the comment made after the CSICOP conference: ‘Perhaps for the time being we should withdraw from tests involving total, synthesised “portraits” of subjects and begin nailing down certainties in connection with trait after trait after trait’ (Deachman 1989).

Applying the semiotic principle: Group 2 — as a projective technique
graphology is weak or redundant: problems of prediction

Why do we need personality tests at all? Is there a general human facility for secrecy and subterfuge in self-presentation that necessitates a reliable (meaning empirically based) assessment of an individual’s behavioral characteristics? Hw is thought to signal the causes of communication difficulties in adolescence and emotional partnerships, but it is in the lucrative area of employment decisions that it jostles for recognition as a projective technique. There it finds no impressive competition, and I suggest it is a fallacy for graphologists to suppose, as many do, that the correlation of good AB hypotheses with the findings of standardly used psychometric tests shows graphology to be scientific or validates its claims. Psychometric tests measure personality and aptitude. Fowler (1991: 43) quotes the BPS in giving personality tests the very low validity coefficient of 0.15, and Blinkhorn (1990: 672) demonstrates that there is little evidence of enduring relationships between the test scores ‘of the most reputable tests [including the Occupational Personality Questionnaire, doyen of personality tests] in the hands of specialists’ and measures of success at work. Testers themselves admit that the validity of all the tests is disappointingly low (Hubbard 1985). Given these doubts, very little inference can be drawn from comparisons between graphological and psychometric results (Fowler 1991; Dixon 1990). Furnham (1988) claims that findings from Eysenck’s (1945) test on extroversion-introversion, neuroticism and psychoticism ‘relate consistently, theoretically and predictably to physiological, psychopathological, cognitive and social variables’. Equally, we could say on the basis of one very successful test (e.g., Zweigenhaft 1973 or Douglas and Sara 1975) that graphological conclusions provide a yardstick for measuring personality. The validity of both techniques is overall equally tenuous, and (Fowler 1991) ‘if the two sets of results correlate, it seems just as probable that both are wrong as both are right’. Although psychologists wishing to test graphology naturally construct psychometry–graphology comparisons, graphology is more fairly tested against observable data.
Graphology’s claim to supersede other diagnostic tools embraces the notion that lack of correlation could prove graphology right and other forms of psychometry wrong. Psychologists hold that because graphology often shows low validity when related to psychometric tests, it is a redundant technique (2b). Why use it if other cheaper, more accurate, efficient measures exist? But high correlations with the results of psychometric tests would make graphology redundant. They would simply confirm the awesome capacity of currently practiced psychometry. The fact is that the practical results of standard personality tests are so weak that firms are turning from personality tests to graphology (Bar-Hillel and Ben-Shakhar 1986: 263). Consultancies claim that even the results of the OPQ require careful judgment in interpretation (Hubbard 1985), so castrating their role in personnel selection and reducing the efficiency of the questionnaires if not making them redundant. Their obvious flaws (misunderstanding of questions, limited choice in questions, self-assessment in answers, and test sophistication) are increasingly remarked. On standard projective techniques Dixon (1990) comments: ‘If it is possible to form judgments of folk from their reflections on the meaning of ink-blots [Rorschach test] there seems no reason in theory why the same should not apply to their hw’. Ink blot interpretations rely on iconicity and rampant symbolism: no fixed link in nature is thought to exist. Figure 1 shows the potential of graphology to back its claim with a causal story. Psychologists appear to accept the low validity of psychometric tests and find no need for further justification of them.

Despite the view that even if there is such a thing as personality there can be no objective agreement on it through testing (Blinkhorn 1990), I believe that graphology has the potential, as a projective technique, to reveal behavioral characteristics. That said, the claim should be modified for current testing to a revelation of the character mode of the writer at the time of writing. Present behavior does not necessarily match behavior under future conditions (Bar-Hillel and Ben-Shakhar 1986: 272–273; Klimoski and Rafaeli 1983; Williams and Stuparich 1990). Therefore regarding 2c (Predictive capacity is low, especially relative to employment decisions), graphologists should not make predictions as to the future success of a particular individual in circumstances about which their knowledge will be incomplete (Molander 1991). Note that ‘predications’ and ‘predictions’, conflated by Goldberg (above), Ben-Shakhar (above), and Furnham throughout his 1988 paper, are not equivalent. Predications are descriptions of present behavior; predictions relate solely to conjectures about future behavior, a distinction well made in Fowler (1991). Graphologists can only offer a systematic assessment (a set of predications) as to the presence or absence of defined characteristics now. To
return briefly to the animal paradigm, it was found that behavior displayed by different species varied even in tightly regulated experimental environments: once the animals were released from conditioning situations and placed in a more natural setting, their behavior patterns were severely disrupted (Mackenzie 1977: 161–162). Sandahl (1991) found that the six graphologists who analyzed one female job applicant’s handwriting achieved a high degree of consistency in description. They disagreed, however, in their predictions of future job performance: three recommended her and three regarded her as unsuitable for the position.

But have not tests validated a priori assumption (iv), that handwriting is a unique, stable feature of an individual even under different conditions? If we then say that prediction of behavior in different circumstances is impossible, are we not saying that there is no link between stable handwriting traits and modes of behavior? This is only true if the handwriting is exactly the same in all its features at the later time (T2) when the ‘personality profile’ seems proved wrong and the prediction falsified as it was at the time the predictions were made (T1). As a fair test against observable data, graphological assessments regarding employment suitability should list precisely the script features, together with their correlated character traits, on which employment has been recommended. This information should be stored, to be reassessed after actual aptitude and performance have been established. The original script features on which judgments were based may have diminished or vanished. Alterations in graphic indicators may be minimal and subtle, but they will be clearly observable in handwriting, and they may be significant, correlating directly with changes in behavior. If re-analysis finds script changes specifically relevant to behavior changes, graphology’s claims are vindicated with regard to those AB correlations. Slyter’s 1981 study would be ideal material for such treatment, since the handwriting samples and assessments were, in fact, stored. My own handwriting provides a relevant example of barely observable changes. It appears in all respects similar to that of ten years ago. But it is smaller and more simplified, and this is of graphological significance. Changes in size and simplification are thought to affect choice of environment and occupation, and aptitude within it. It would surprise no graphologist that I have moved in that time from (roughly) right-brain and imaginative activity to left-brain concentration and academic work (Wellingham-Jones 1988). Graphological assessments regarding employment suitability need to list precisely at T1 the handwriting traits on which employment has been recommended. These are predications or descriptions of character. To reassess them at T2 after aptitude has been established is to redescribe the newly emerged character indicated by a later set of script features.

Williams (1990) suggests two factors which can affect measured perfor-
mance: firstly, changes as the result of unforeseen events or intellectual growth, and secondly, changes in the global or economic environment in which the personality profile as defined at T1 might be expected to perform. Naftali (1986) imputes hw changes to fluctuations in the organic or biochemical balance, which is also responsible for behavior. We live in a probabilistic universe (Williams and Stuparich 1990). This sets limits for graphology as for all other predictive fields, including to a lesser extent the organic sciences. Because handwriting is a physical expressive field which aspires to the status of indexical sign system, graphology may possibly avoid some of the problems of economics, sociology, psychology, and other areas of investigation where speculation, feelings and subjectivity, and other relativistic data are the major sources of information. But it cannot free itself from the real organic change which behavior — character — is believed to reflect.

**Applying the semiotic principle: Group 3 — graphology requires no skill**

We can dispose at once of the notion that graphologists simply read the information given in hw samples (Bar-Hillel and Ben-Shakhar 1986; Ben-Shakhar et al. 1986). This charge has been dropped by serious researchers following the conclusion of Rafaeli and Klimoski (1983) that 'script content has little effect on graphologists' assessments'.

The criticism that graphological judgements are no better than chance is ambiguous. It can mean simply that test results show coefficients of correlation no better than chance would throw up (0.1–0.25). In the case of predictive validity this is probably true, as it is for all other projective techniques. Alternatively, Group 3 gives rise to an irritating problem which the semiotic principle illuminates. Critics suggest that many non-graphologists receive and interpret impressions from hw and can match trained graphologists in accuracy of judgment. Such guesswork has nothing to do with scientific method; the semiosis at work is iconicity. Resemblances between the look of the hw and the state of affairs extrapolated from it are almost certainly culturally perspectival, and may have many acceptable but contradictory interpretations. Teillard’s love letter written in Jukajire tribal picture language is interpretable only insofar as we recognize its images (Teillard 1983: 186). Few Westerners would grasp that the big tree and little tree are lovers, that the fat tree is a female rival, or that the intestine-like protuberance from one of them signifies a young girl’s dream. Iconicity often accounts for unsatisfactory test conclusions. In one test graphologists identified genuine suicide notes and picked out the fakes, but rather than positing specific correlations between
hw traits and suicidal behavior they ‘subconsciously’ identified signs (Frederick 1968). Only indexical signs are adequate as an explanation of graphological judgments, and if graphologists want to be considered scientific practitioners they must use only such signs. Graphological invitations to note ‘axe’ or ‘knife blade’ shapes in hw, or i dots shaped like the 6 in ‘666 — the devil’s sign’ (Sassi 1987), are the hallmarks of a showman, not a scientist.

Graphology practiced as an iconic sign system would turn the subject into ‘cold reading’, a technique with no methodology or empirical base, but instead reliant on the intuitive (or as is sometimes claimed, the psychic or supernatural) powers of the practitioner. Historically, graphology has no roots in this kind of technique. The seventeenth-century conclusions of the physician Baldi, like those of L’Abbaye Michon, were not based on iconicity, but were inductively derived by matching observations of hw traits to observations of character. Induction is not considered to validate hypotheses, but is legitimately used in constructing them. The inference from a hw variable to a character mode is allowable only after experimental demonstration. Lockowandt (1986) grumbles that ‘There is no sensible reason for the lack of academic interest in this subject. It is absurd to cite its somewhat dubious past; we cannot raise alchemy as an objection to chemistry’. It is sadly true that graphology has a more dubious present than past to cite: Baldi and Michon, meticulous empiricists in their day, would grieve at iconic interpretations which have no basis in systematic observation and are often introduced for effect. Serious graphologists attend to such features no more than they do content, and only vanity, greed, or ignorance of the implications of their actions drives them to demonstrate at psychic festivals or participate in tests of the paranormal.

Conclusion

Realists believe that truth lies not in a web of belief we weave for ourselves (though this was Peirce’s conviction — CP 4.531 et seq.), but ‘out there’ in the world, independent of our perception of it. Scientific understanding of the physical construction of the world, including human beings as natural objects, has transformed Western science and all branches of ethology. The rapid advance of Western science speaks loudly for its methodology. But it would be to give an incomplete view of how scientists approach their work not to mention limitations. We should not overestimate the status of scientific conclusions, however well corroborated. Discovering how things are is inevitably perspectival: we have a
specifically human viewpoint in that we only have our own meager physical faculties and their technical brainchild to explore the world of which we are a part. The content captured in a model is thus a function of the model and of the processes which evaluate it. Graphologists who hold that their study can never be truly scientific are not alone. Many social scientists, psychologists, and even medical doctors share this view. They believe their theories are often the result of communal discussion, and that their truth is grounded in intersubjective opinion. We may yet confirm a model of a criminal which openly depends on our agreement in discourse rather than on some glimpsed, ‘verified’ reality.

But intersubjective agreement could talk fairies into existence. While graphology ultimately has no compulsion to seek the status of exact science, I believe it has an empirical basis which it should, in its best interest, exploit. I have applied strict semiotic principles to show its potential as an indexical sign system, and I think it should press its claim that its semiosis is real not because our opinions converge but because of how things, in fact, are. My suggestion for a more satisfactory research paradigm cleaves to this assumption. The assumed link between hw and character modes is backed by our best current understanding of physical mechanisms, especially as they relate to motor control and biochemistry. Exploratory studies of CB and CA correlations (e.g., Wing and Baddeley 1978; Wing et al. 1991) should be followed up with BA hypothesis testing as detailed above to ensure that a false correlation has not been posited. Conversely, graphological hypotheses concerning mental disorders should, where possible, be tested clinically. In Douglas and Sara (1975) particular hw traits believed to signify schizophrenic behavior (BA) were found to correlate precisely with states of the disease (CA, CB). Two-way testing of this kind, if confirmed, signifies an undeniable correlation: this specific B can reliably be taken to be a sign of this specific A. If graphology can achieve this degree of corroboratation it will have surpassed that of much normal science. Such results should not only be published but publicized, with the term ‘graphology’ given its share of the success. It will, in the narrow sense, have vindicated the claims made in the opening sentences of this paper.

Notes

1. Revised and extended version of a paper read at the Second Symposium of Graphological Research at King’s College Cambridge, 12 August 1989. I am grateful to Mr. David Houghton, of the University of East Anglia, for helpful comments and stimulating skepticism during the writing of the original paper, and to my daughter


—(1991). Research papers relating to the validity and reliability of graphology. Mailing address: The British Institute of Graphologists, 4th Floor, Bell Court House, 11 Blomfield St., London EC2M 7AY.

Harvey, O. L. (1934). The measurement of handwriting considered as a form of expressive movement. Character and Personality 2, 310–321.


Amanda for input from her study of experimental psychology at the University of Cambridge.


References


